The Effect of Free Flight on Air Traffic Controller Mental Workload, Monitoring and System Performance

Brian G. Hilburn, Marcel W. P. Bakker, Wouter D. Pekela
National Aerospace Laboratory (NLR), Amsterdam
Raja Parasuraman
Catholic University of America, Washington, DC

Abstract

Under mature free Flight (FF), aircraft outside of terminal areas would generally be free to fly their preferred routes, and self-separate, with minimal intervention from air traffic control (ATC). From an ATC perspective, FF could raise a number of human performance problems (including workload extremes, passive monitoring demands, and difficulties in reverting to manual control). This article describes a moderate-fidelity ATC simulation recently carried out at the NLR, in which *Controlled Flight* conditions (analogous to current-day operations) were compared to *Free Flight* for the en-route environment. The simulation specifically manipulated *Intent Sharing* under FF— that is, whether aircraft provided advance notice of their intended manoeuvres. Results showed workload benefits of FF (especially under high traffic). Intent information seemed to increase controllers' acceptance of FF, but had no clear effect on either workload or monitoring performance. Finally, HMI considerations emerged as important ones for future work in this area.

1. Introduction

Free Flight (FF) has been proposed as a way to both handle ever-increasing air traffic demands, and to provide economic benefits to airspace users. Although FF has thus far been defined only at a high level (RTCA, 1995), research into FF concepts (e.g., direct routing) is proceeding on both sides of the North Atlantic. According to a vision of mature FF, aircraft outside of terminal areas would generally be free to fly user-preferred routes, and modify their trajectories en route, with minimal intervention by air traffic control (ATC). Although the advent of FF assumes certain enabling technologies (e.g., ADS-B capability, and conflict probe tools), FF would represent as much an operational, as a technological, evolution.

Under likely near-term FF scenarios, the Air Traffic Controller (ATCo) would continue to play an important (albeit new) role in ATC (Hanson, 1997), especially in the face of unpredictable aircraft behaviour. Rather than strategically controlling air traffic, the "controller" of the future might well fill the role of a "Separation Assurance Monitor," intervening tactically only when losses of separation are imminent. This new role raises a number of potential human performance problems, including:

- Workload extremes (either underload or overload);
- Passive monitoring demands (i.e., vigilance);
- Difficulty in reverting to conventional control (e.g., emergencies);
- Intent ambiguity in joint air / ground displays and algorithms.

A recent experiment by Endsley, Mogford & Stein (1997) assessed the effect of FF-like scenarios on ATCo situation awareness and mental workload. Workload in their study, however, was only assessed in terms of self-reported subjective workload. This paper reviews an exploratory experiment recently conducted at NLR into the effects of similar FF traffic scenarios on ATCo workload, monitoring performance, and ability to anticipate non-nominal situations. This was done by assessing the performance of currently-active controllers under both conventional (i.e., controlled) and free flight conditions, using the same en route airspace. Two free flight conditions were evaluated: one in which aircraft shared their intentions with ATC before manoeuvring, and one in which aircraft manoeuvred without notifying ATC. In addition to subjective workload ratings, the current study also collected objective (pupil diameter) measures of mental workload.

2. Method

2.1. Air Traffic Controllers Test subjects were ten United Kingdom Royal Air Force (RAF) military controllers, drawn from both the Glasgow and London regions. Of these, all but two were currently active controllers. The final two controllers had recently been retired from the RAF. Age ranged from 30 to 40 (mean = 35.5 years), and years of active controlling experience ranged from 6 to 22 (mean = 11.9).

2.2. ATC Task The experiment was based on a simulation of the Maastricht-Brussels en route airspace, in which controllers normally handle traffic along several intersecting paths. The experimental airspace is depicted in Figure 1. Four traffic samples were created, all based on the same master traffic scenario. The master traffic scenario was carefully created and checked for realism (e.g., callsigns and SSR codes) by a subject matter expert. Slight modifications to the master traffic scenario yielded four highly similar (though non-identical) traffic samples, each 75 minutes in length. Traffic density was varied within each session to provide realistically extreme levels of traffic load. Traffic samples were checked and pre-tested for realism and traffic load equivalence (in terms of flight entry rate).

Whereas aircraft in conventional traffic samples were scripted to manoeuvre along the air route and beacon system, FF traffic samples followed a direct routing structure, as specified by a set of 32 Trajectory Change Points (TCPs) around the perimeter of the sector. These TCPs limited the number of points through which an aircraft might enter/exit the sector. Although FF conditions provided no flightplan as such, the display of entry and exit TCPs was under the control of the ATCo. Under FF conditions, aircraft appeared to generally track direct routes between these entry and exit TCPs, and manoeuvre only as needed to self-separate. Thus the FF traffic scenario simulated two key elements of a mature FF environment: (1) direct routing, and (2) self-separation. Figures 1 and 2 depict the display differences between CF and FF airspace, as well as between low and high traffic densities.

Under all flight conditions, ATCos were responsible for accepting and handing-off aircraft at sector boundaries. Under CF, controllers had to issue commands by Radio

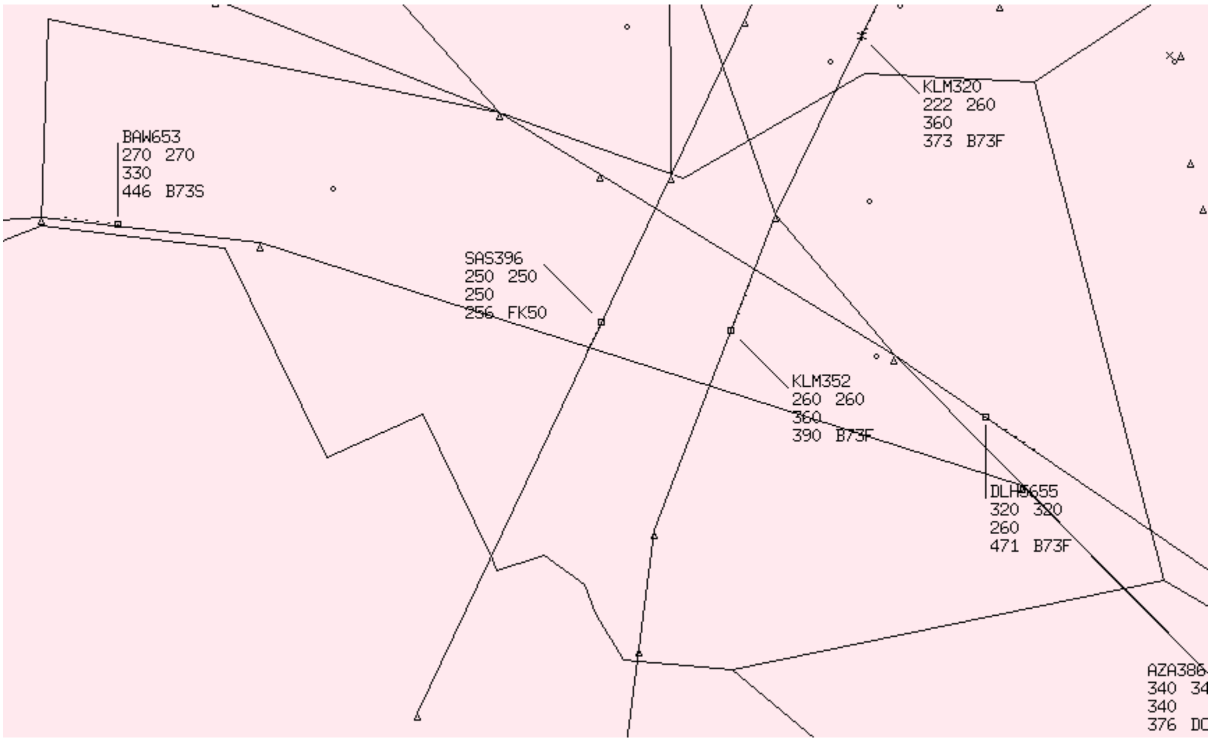

Figure 1. Maastricht-Brussels en-route airspace (CF condition, low traffic).

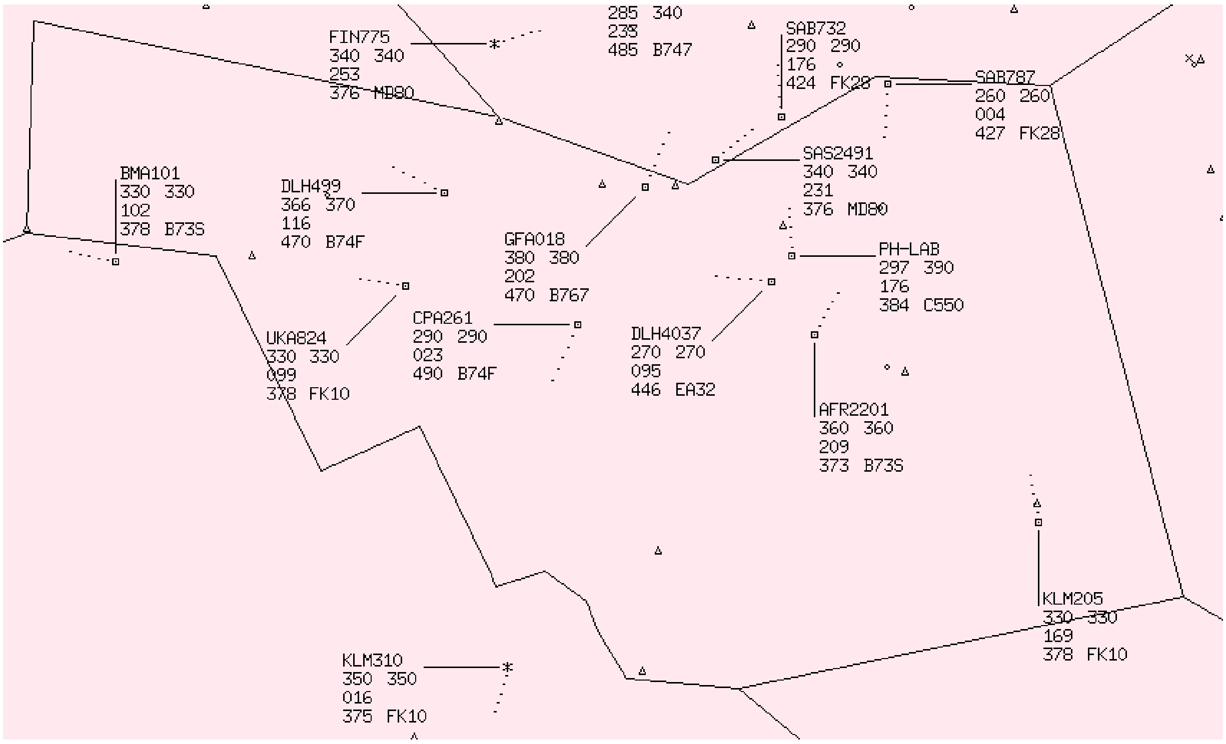

Figure 2. Maastricht-Brussels en-route airspace (FF condition, high traffic).

Telephony (RT). It was recognised that permitting controllers to exercise their preferred control strategies might deprive us of any data under FF conditions—that is, ATCos might be reluctant to actually permit FF. As a result, ATCos were instructed to intervene in the FF traffic pattern only in the case of an STCA warning (several were scripted per FF session), when tactical avoidance was required. To permit an assessment of the control strategies controllers would have preferred to use on the free-flight samples (as well as to control somewhat for the level of speech— which could have introduced measurement artefacts in the pupil diameter data) between controlled and free-flight sessions, a verbal call-out procedure was used whereby subjects identified aircraft pairs according to a three-point separation criticality scale, as follows:

Level 1 Alert—

- I WOULD permit this situation under *controlled flight* conditions;
- Corrective action might be required in the future;
- I would continue to monitor this situation.

Level 2 Alert—

- I would NOT permit this situation under *controlled flight* conditions;
- Corrective action MIGHT be required in the future;

Level 3 Alert—

- Corrective action WILL PROBABLY be required.
- Loss of separation is imminent.

The operational scenario for this experiment included some simplifying assumptions regarding the use of sector-specific Special Use Airspace (SUA). It also included a set of "Rules of the Air," known as *Extended Flight Rules* (EFRs), that were intended to dictate how FF aircraft should self-separate, under conditions of minimal (or no) ground intervention. They had to do so both comprehensively (i.e., for all possible traffic encounters) and unambiguously (i.e., each party had to have a clear understanding of the responsibilities of all aircraft). Further, to expedite training, it was decided that the number of extended flight rules had to be kept to an absolute minimum. These EFRs were based on a synthesis of (1) ICAO extended VFR overtaking rules (e.g., "overtake on other ship's starboard side"), (2) Altitude for Direction of Flight (i.e., "east is odd Flight Level"), and (3) Phase-of-Flight priorities, as specified under Eurocontrol's FREER project (Duong & Floc'hic, 1996).

This experiment was carried out using the NLR ATC Research Simulator (NARSIM), which provided for exact scripted control over the on-screen appearance and behaviour of aircraft. Simulated aircraft were operated by a team of pseudopilots, either under the control of the ATCo test subject (under controlled flight conditions) or in accordance with session scripts (under free flight conditions). Under free flight conditions, the appearance of co-ordinated airborne self-separation was simulated through the use of a Medium Term Conflict Alert (MTCA) that permitted the pseudopilot team to avoid or permit conflicts (as session scripts dictated). The ATC plan view display was presented on a Sony 2,000 x 2,000 pixel screen. Although interface modifications to the baseline NARSIM were minimised, free flight conditions did require the following display changes:

- Flight Data Block (FDB) presentation of both [1] ATCo-commanded, and [2] Aircraft-selected parameters (e.g., heading, speed);
- Suppressed display of routes and beacons under free flight sessions.

The format and appearance of the flight data block labels are shown in figure 3, for each of the three flight conditions. When ATCos assigned either a heading or altitude under FF (a non-nominal intervention) this was reflected in a third column within the data label. In figure 3, for instance, the ATCo has assigned the aircraft to a heading of 180, and has stopped its descent short (FL 300) of the aircraft's self-selected bottom of descent, FL 290.

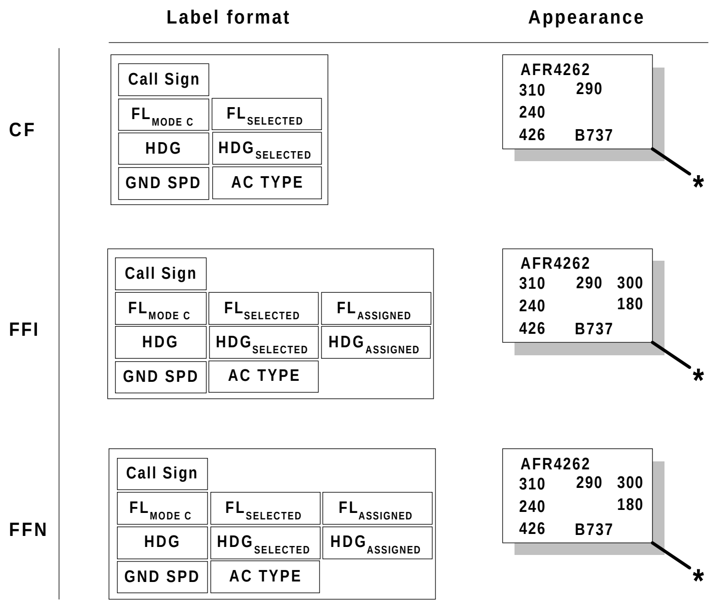

Figure 3. Flight data label format and appearance, by flight condition.

<u>2.3. Experimental Design and Procedure</u> This experiment manipulated the following two factors in a repeated measures design: Flight Condition (3 levels), and Traffic load (2 levels). ATCos were provided familiarisation materials (regarding the task display and experimental protocol) in advance of their on-site participation. After a half-day of on-site familiarisation and training, each ATCo completed three 75-minute experimental sessions. Traffic Load was varied over the course of each session, with periods of Low and High traffic (counterbalanced within each session) averaging roughly 10 and 17 aircraft under simultaneous control. The three levels of Flight Condition were defined (and their order randomised across ATCos), as follows:

- *Controlled flight (CF)*— aircraft navigated according to standard route structure (unless instructed otherwise by controller), and manoeuvred only in response to controller-issued clearances;
- *FF <u>with</u> Intent Sharing (FFI)*— Route structure was neither displayed nor used, flightplans provided only sector entry/exit points, and aircraft shared their intentions with ATC before initiating any action;
- *FF <u>without</u> Intent Sharing* (FFN)— As above, although aircraft actions were not pre-announced to ATC.

Data from this study included a number of controller workload metrics, as well as system monitoring performance. Workload measures were of two types: Objective (pupil diameter) and subjective (the Rating Scale for Mental Effort, or RSME (Zijlstra & van Doorn, 1985)). Previous experience has shown these measures to be sensitive and reliable indicators of workload in simulated ATC tasks (Hilburn, Jorna and Parasuraman, 1995).

Pupil diameter data were collected with the Observer® eye tracking system, once per gaze fixation, with a theoretical resolution of .04mm. The RSME subjective workload scale is a simple paper-and-pencil instrument that requires subjects to indicate workload, on a continuous unidimensional scale. Controllers were instructed at several points throughout each session to rate their current workload using the RSME instrument. As appropriate, statistical analyses for all measures were carried out through univariate Analyses of Variance (ANOVAs).

Data analysis focused on the following research questions:

- What is the effect of FF traffic patterns on controller mental workload?
- Does the potential loss of aircraft intent information under FF impact controllers' response time to non-nominal events (e.g., STCA warnings)?
- Does this loss of intent information seem to degrade controllers' ability to anticipate critical events (e.g., losses of separation)?
- What are controllers' subjective impressions of a FF-like operational scenario?

3. Results

3.1. Controller workload

Again, controller workload was assessed using both objective (pupil diameter) and subjective (self-report) measures, and these will be discussed in turn.

3.1.1. Pupil diameter

Increases in visual workload (as opposed to, say, memory load) are generally associated with small but measurable increases in pupil diameter. Indeed, a statistically significant difference ($p<.001$) was found between pupil diameter under low and high traffic—pupil diameter was seven percent (7%) higher under high traffic than under low traffic. The trend depicted in figure 4 shows that indicated workload was lower under the two FF conditions than under Controlled Flight. This was especially true under high traffic conditions.

3.1.2. RSME Subjective Mental Workload

Controllers reported significantly higher workload under high traffic conditions, ($p<.0001$). No main effect of control condition was found on RSME scores. A significant interaction was found between traffic level and control condition, ($p<.05$). This interaction is depicted in figure 5. A post hoc Newman-Keuls test revealed that, under high traffic, controllers felt significantly more workload under controlled flight than they did under uninformed free flight (FFN), ($p<.05$).

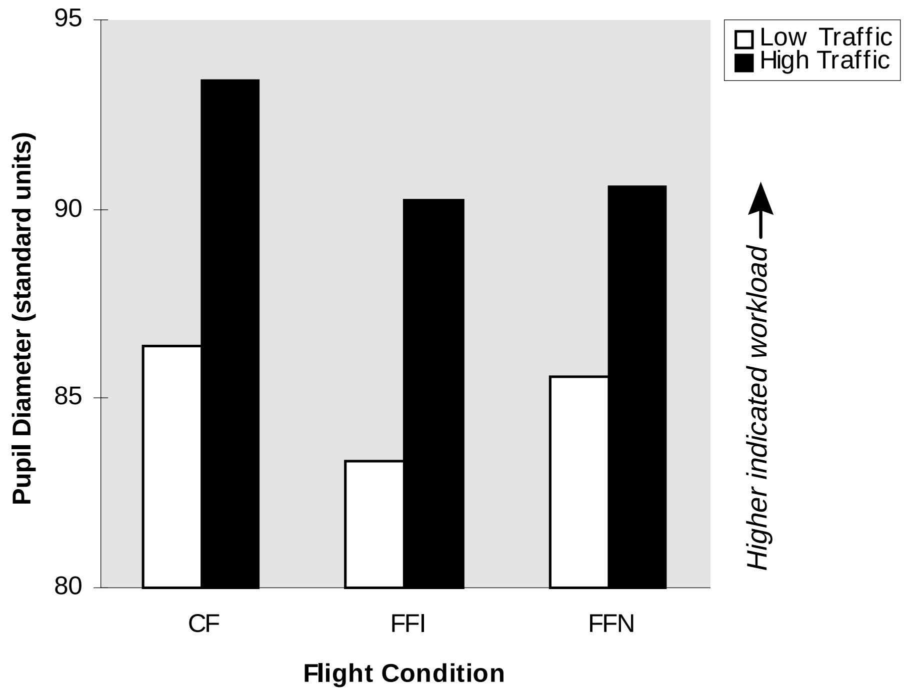

Figure 4. Pupil diameter, by Flight Condition and Traffic Load.

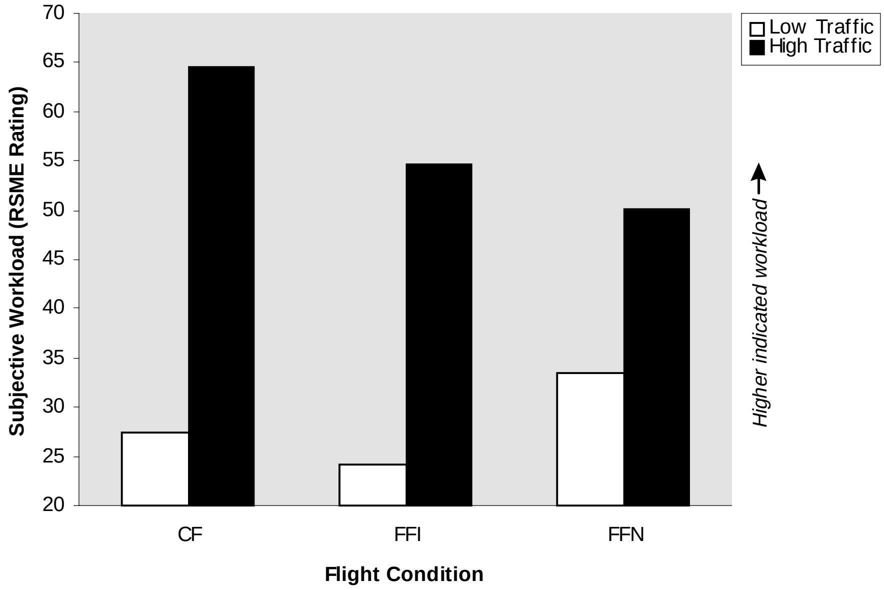

Figure 5. Self-reported workload, by Flight Condition and Traffic Load.

3.2. Monitoring Performance and Traffic Awareness

3.2.1. Response time to STCAs

Mean response times (from STCA onset to issuance of a corrective clearance) were calculated. Mean response times were lower for the CF than for either the FFI or FFN conditions, at 8.6, 10.0 and 9,9 seconds, respectively. FF traffic scenarios were scripted to provide a fixed number of situations (e.g., STCAs). This was obviously not the case under CF conditions, and as a result the number of STCAs differed dramatically between CF and FF conditions. Because of this difference (and the corresponding difference in standard deviations), statistical analysis of the response data (CF vs FF) is not appropriate. Comparing the two FF conditions, however, showed no significant difference in response times (10.0 secs. versus 9.9 secs.) between the informed (FFI) and uninformed (FFN) FF conditions—indeed, average response time to STCAs was slightly lower for the uninformed FF condition.

3.2.2. Conflict Prediction Accuracy under FF

Prediction accuracy was defined as the proportion of all STCAs for which a given controller had reported a potential conflict situation (according to the three-point severity scale), irrespective of the number of "false alarms" (i.e., situations in which a reported conflict did not result in an STCA)[1]. These data were available only under the FF conditions (since, under CF conditions, they were free to proactively control the traffic).

Controllers generally did not anticipate all STCA situations; Detection rates ranged from 0% to 100%. A statistically significant difference was found between prediction accuracy under low and high traffic, with averages of 88.0% and 33.2% under low and high traffic, respectively, ($p<.02$). Under FFI, controllers correctly predicted 64.9% of all STCA situations, whereas under FFN they correctly predicted only 53.3%. This difference failed to reach statistical significance. A trend toward lower prediction accuracy for Uninformed FF appeared only under low traffic.

3.3. Intervention and Control Strategy Differences

All ATCo inputs and system interactions were logged during test sessions. Among the parameters logged was the occurrence of flightplan information requests made by the controller. The pattern of such information requests is shown in figure 7, by both Flight Condition and Traffic Level. Consistent with Endsley, Mogford and Stein (1997), who noted that FF might increase controllers' tendency to query aircraft, these data show that more flightplan information requests were made under FF than under CF conditions. The fact that number of queries decreased with traffic might simply be an indication of spare capacity-- ATCos might have tended to query only as time permitted. [2] Comparing the number of flightplan queries under the two FF conditions,

---

[1] No pattern was discernible in controllers' false alarms.

[2] Notice that, even under FF, ATCos remained responsible for hand-off and acceptance of aircraft, so the task was not one of entirely passive monitoring.

it is interesting that fewer queries were made under the FFN condition, in which aircraft were not sharing their intentions with ATC.

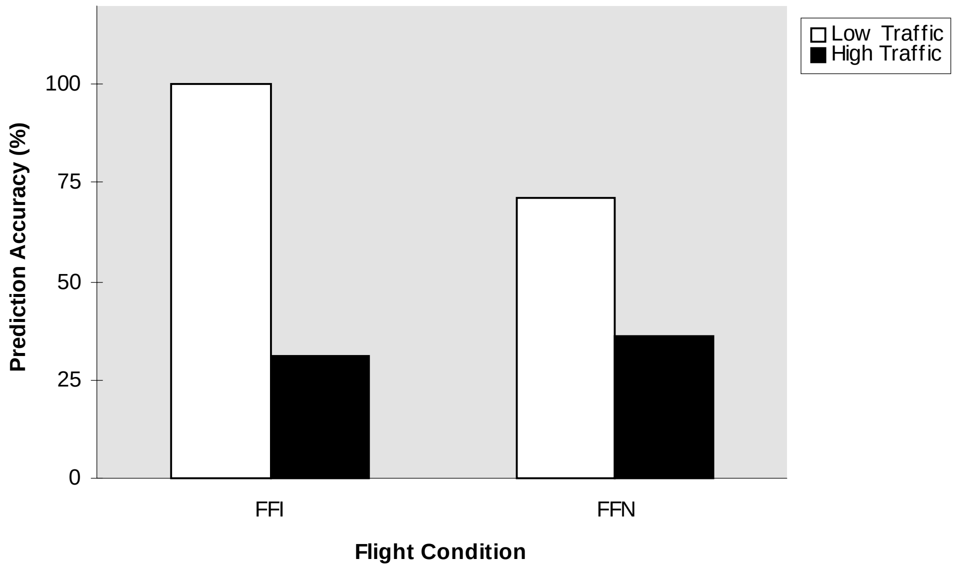

Figure 6. STCA prediction accuracy, by Flight Condition and Traffic Load.

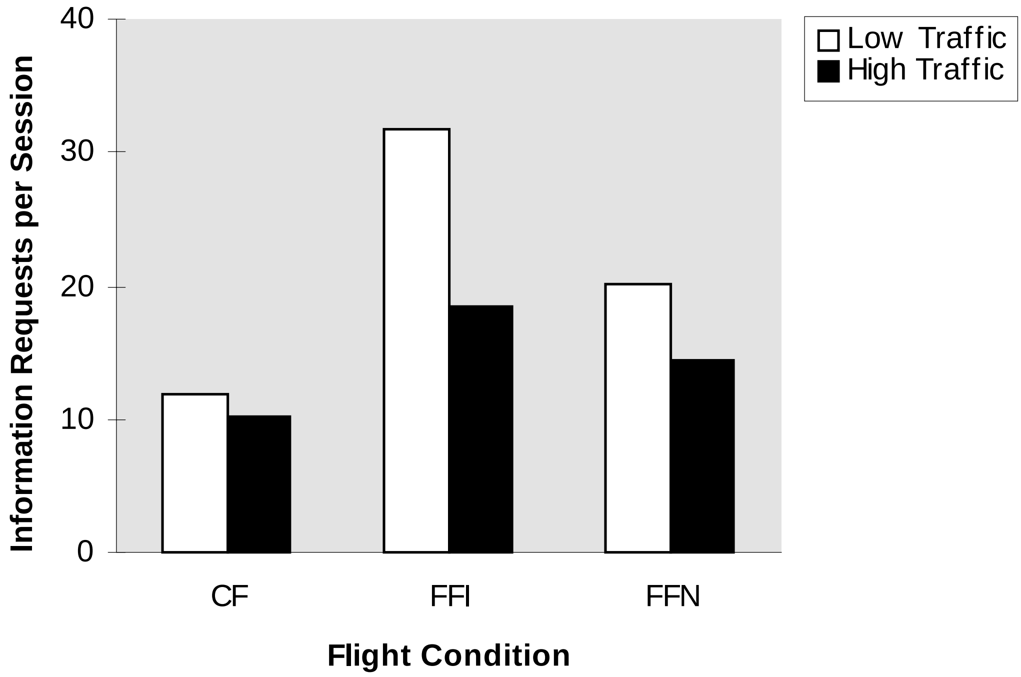

Figure 7. Average number of flightplan information requests per session, by Flight Condition and Traffic Load.

3.4. General Observations and Feedback

Because this study was intended as an exploratory investigation of FF concepts, the experimental team sought to elicit as much subjective feedback as possible from participants. Some general observations include the following:

- Controllers generally found FF surprisingly easy, and reported that workload was much lower than they had anticipated;
- They also felt strongly that aircraft intentions should always be available to the controller;
- Controllers generally felt that the Rules of the Road used for this study were clear, and facilitated detection of unusual situations;
- Most felt that conflict detection was more difficult under FF;
- Opinion was evenly split on whether STCA provided an adequate safety net function;
- Several controllers expressed concern that controllers under FF would be forced to over-rely on STCA, thereby depriving themselves of sufficient time and control options to resolve situations;
- Several controllers also volunteered that if aircraft had been free to communicate their intent to both ATC and to other aircraft, ATC could become safer and easier;
- Most controllers reported on shortcomings of the PVD interface. The need for label de-cluttering, ICAO destination designators (in flight data blocks) and velocity trend vectors were issues most mentioned.

4. Discussion

Trends in both objective and subjective workload measures suggest that FF can reduce workload, relative to CF conditions. Under low traffic density, the indicated workload reductions were greater for FFI than FFN—that is, shared intent information reduced controllers' indicated workload. Under high traffic, however, there was no pattern to suggest that shared intent information between air and ground reduced the controllers workload.

Given the large projected increases in air traffic, proponents of mature FF would probably be more interested in the high traffic density condition. Under high traffic, the indicated workload benefits of FF were more apparent. The objective and subjective measures were in essential agreement with one another, which in other studies of new ATM concepts has not always been the case (Hilburn, Jorna, and Parasuraman, 1995).

Monitoring data revealed that having manoeuvre intent information (FFI versus FFN) did not reduce response time to short term conflict alerts. Further, STCA prediction accuracy was nearly identical under high traffic conditions, regardless of whether intent information had been provided.

Consistent with the study of Endsley, Mogford and Stein (1997), controllers queried the system more under FF than under controlled flight conditions. Perhaps surprisingly, though, under FF controllers tended to query more when maneuver intent information was provided.

Summing up results, these data suggest the following:

- FF seems to reduce controller workload, especially under high traffic;
- Under FF, lack of manoeuvre intent information does not worsen controllers' ability to predict traffic changes;
- Intent information in general seems to benefit controllers' acceptance more than their workload;
- Controllers query the system more under FF, although are more likely to do so if manoeuvre intent information is provided;
- Controller acceptance of the FF concept might be fairly high;
- HMI display considerations will have to be further addressed in developing controller tools for FF.

The results of this experiment suggest that the potential human performance costs (e.g., mental workload increases) of FF might be smaller than those demonstrated by Endsley, Mogford and Stein (1997). In explaining this discrepancy, it might be instructive to consider two major differences between the two experimental protocols: First, the current study employed military ATCos, who (because of various operational differences from their civilian counterparts) might represent a more appropriate population with which to test fundamental aspects of FF (cf., Hilburn & Parasuraman, 1997). Second, the current study employed a fuller battery of workload measures. It is hoped that these measures together can provide a fuller picture of the influence that FF might have on the performance of the future ATCo.

The current experiment was intended as an exploratory analysis, to gain familiarity with some of the most salient aspects of possible FF operations (i.e., direct routing and self separation), and to gain an understanding of how these might impact human and system performance. As a result, interface changes were kept to a minimum. It is recognised that a mature FF environment would likely bring with it requirements for vastly redesigned ATC displays and control algorithms. This point is emerging from data on transient workload. Preliminary data analysis suggests that the workload benefits of manoeuvre intent sharing appear only <u>transiently</u>— namely, during the occasional non-nominal situation (such as during separation alerts). If this is so, it raises the possibility that advanced displays of aircraft manoeuvre intent might prove beneficial during such non-nominal situations. For instance, if ATCos were provided real-time displays that enabled them to verify that aircraft were co-ordinating joint evasive manoeuvres, would this reduce ATCos' transient workload? At least as importantly, would it allow them to better assess whether to intervene in the traffic pattern? These issues are to be explored further in an upcoming experient.

6. References

Duong, V. & Floc'hic, L. (1996). *FREER-1 Requirement Document version* 2.0. Bretigny, France: EUROCONTROL Experimental Centre.

Endsley, M.R., Mogford, R.H. & Stein, E. (1997). Effect of free flight conditions on controller performance, workload and situation awareness: a preliminary investigation of changes in locus of control using existing technology. Draft FAA technical report. April, 1997.

Hanson, T. (1997). Bordering on applied chaos. Air Traffic Management. vol. 6 (4), 18,20-24.

Hilburn, B. & Parasuraman, R. (1997). Military Controllers Provide an Especially Appropriate Population for Evaluation of Free Flight and Other Advanced ATM Concepts. Poster presented at the CEAS 10th European Aerospace Conference on Free Flight. 20-21, October. Amsterdam, The Netherlands: Confederation of European Aerospace Societies.

Hilburn, B., Jorna, P.G.A.M. & Parasuraman, R (1995). The effect of advanced ATC strategic decision aiding automation on mental workload and monitoring performance: an empirical investigation in simulated Dutch airspace. In *Proceedings of the Eighth International Symposium on Aviation Psychology*, April 1995, Columbus, Ohio USA: Ohio State University.

Zijlstra, F.R.H & Doorn, L. van (1985*). The construction of a scale to measure subjective effort*. Technical Report, Delft University of Technology, Department of Philosophy and Social Sciences.

7. Glossary

ADS-B	Automatic Dependent Surveillance - B
ATC	Air Traffic Control
ATCo	Air Traffic Controller
BSMI	Rating Scale for Mental Effort
CF	Controlled Flight (also an experimental condition)
EFR	Extended Flight Rules
FDB	Flight Data Block
FF	Free Flight
FFI	FF with information sharing (experimental condition)
FFN	FF with no information sharing (experimental condition)
FL	Flight Level
ICAO	International Civil Aviation Organisation
MTCA	Medium Term Conflict Alert
NARSIM	NLR ATC Research Simulator
NLR	National Aerospace Laboratory of the Netherlands
PVD	Plan View Display
RSME	Rating Scale for Mental Effort
SSR	Secondary Surveillance Radar
STCA	Short Term Conflict Alert
SUA	Special Use AIrspace
TCP	Trajectory Change Point
TLX	Task Load Index
VFR	Visual Flight Rules